22DCE006-Probin Bhagchandani Practical-1

```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('/content/customer_shopping_data.csv')
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 99457 entries, 0 to 99456
Data columns (total 10 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   invoice_no      99457 non-null  object
 1   customer_id     99457 non-null  object
 2   gender          99457 non-null  object
 3   age             99457 non-null  int64
 4   category        99457 non-null  object
 5   quantity        99457 non-null  int64
 6   price           99457 non-null  float64
 7   payment_method  99457 non-null  object
 8   invoice_date    99457 non-null  object
 9   shopping_mall   99457 non-null  object
dtypes: float64(1), int64(2), object(7)
memory usage: 7.6+ MB
None
```

```python
import numpy as np
blank_arr = np.zeros((3))
print("Blank array (zeros):\n", blank_arr)

predef_data = [2,4,6,8,10]
predef_arr = np.array(predef_data)
print("Array with predefined data:\n", predef_arr)

patt_arr = np.zeros((3,3), dtype=int)
patt_arr[1::3] = 7
print("Specific pattern array:\n", patt_arr)
```

```
Blank array (zeros):
 [0. 0. 0.]
Array with predefined data:
 [ 2  4  6  8 10]
Specific pattern array:
 [[0 0 0]
 [7 7 7]
 [0 0 0]]
```

```python
import numpy as np
a = np.arange(10)
print("original array",a)
s = slice(2,7,2)
print("slice function",a[s])

b=np.array(a-2)
print("updated array",b)
```

```
original array [0 1 2 3 4 5 6 7 8 9]
slice function [2 4 6]
updated array [-2 -1  0  1  2  3  4  5  6  7]
```

```python
arr = np.arange(12)

a = arr.reshape(3,4)

print('Original array is:')
print(a)

print('Modified array is:')

for x in np.nditer(a):
    print(x)
```

```
Original array is:
[[ 0  1  2  3]
 [ 4  5  6  7]
 [ 8  9 10 11]]
Modified array is:
```

```
          0
          1
          2
          3
          4
          5
          6
          7
          8
          9
          10
          11
```

```python
import numpy as np
mydata = np.loadtxt("textfile.txt", dtype=int)
print(mydata)
```

```
[ 1  2  3  4  5  6  7  8  9 10 11 12]
```

```python
import pandas as pd
df=pd.read_csv('customer_shopping_data.csv')
```

```python
df5=df.iloc[0:25]
df5
```

	invoice_no	customer_id	gender	age	category	quantity	price	payment_method
0	I138884	C241288	Female	28	Clothing	5	1500.40	Credit Card
1	I317333	C111565	Male	21	Shoes	3	1800.51	Debit Card
2	I127801	C266599	Male	20	Clothing	1	300.08	Cash
3	I173702	C988172	Female	66	Shoes	5	3000.85	Credit Card
4	I337046	C189076	Female	53	Books	4	60.60	Cash
5	I227836	C657758	Female	28	Clothing	5	1500.40	Credit Card
6	I121056	C151197	Female	49	Cosmetics	1	40.66	Cash
7	I293112	C176086	Female	32	Clothing	2	600.16	Credit Card
8	I293455	C159642	Male	69	Clothing	3	900.24	Credit Card
9	I326945	C283361	Female	60	Clothing	2	600.16	Credit Card
10	I306368	C240286	Female	36	Food & Beverage	2	10.46	Cash
11	I139207	C191708	Female	29	Books	1	15.15	Credit Card
12	I640508	C225330	Female	67	Toys	4	143.36	Debit Card
13	I179802	C312861	Male	25	Clothing	2	600.16	Cash
14	I336189	C555402	Female	67	Clothing	2	600.16	Credit Card
15	I688768	C362288	Male	24	Shoes	5	3000.85	Credit Card
16	I294687	C300786	Male	65	Books	2	30.30	Debit Card
17	I195744	C330667	Female	42	Food & Beverage	3	15.69	Credit Card
18	I993048	C218149	Female	46	Clothing	2	600.16	Cash
19	I992454	C196845	Male	24	Toys	4	143.36	Cash
20	I183746	C220180	Male	23	Clothing	1	300.08	Credit Card
21	I412481	C125696	Female	27	Food & Beverage	1	5.23	Cash

```python
df.to_csv('customer_data.csv')
```

```python
for i,j in df.iloc[:3].iterrows():
    print(i, j)
    print()
```

```
0 invoice_no          I138884
  customer_id         C241288
  gender               Female
  age                      28
  category           Clothing
  quantity                  5
```

```
price                   1500.4
payment_method      Credit Card
invoice_date          5/8/2022
shopping_mall           Kanyon
Name: 0, dtype: object

1 invoice_no             I317333
customer_id             C111565
gender                     Male
age                          21
category                  Shoes
quantity                      3
price                   1800.51
payment_method       Debit Card
invoice_date         12/12/2021
shopping_mall     Forum Istanbul
Name: 1, dtype: object

2 invoice_no             I127801
customer_id             C266599
gender                     Male
age                          20
category               Clothing
quantity                      1
price                    300.08
payment_method             Cash
invoice_date          9/11/2021
shopping_mall          Metrocity
Name: 2, dtype: object
```

```python
cln=df.iloc[:3]
for i in cln:
  print(cln)
```

```
    payment_method invoice_date   shopping_mall
0     Credit Card     5/8/2022          Kanyon
1      Debit Card   12/12/2021  Forum Istanbul
2            Cash    9/11/2021       Metrocity
  invoice_no customer_id  gender  age  category  quantity     price  \
0    I138884     C241288  Female   28  Clothing         5   1500.40
1    I317333     C111565    Male   21     Shoes         3   1800.51
2    I127801     C266599    Male   20  Clothing         1    300.08

    payment_method invoice_date   shopping_mall
0     Credit Card     5/8/2022          Kanyon
1      Debit Card   12/12/2021  Forum Istanbul
2            Cash    9/11/2021       Metrocity
  invoice_no customer_id  gender  age  category  quantity     price  \
0    I138884     C241288  Female   28  Clothing         5   1500.40
1    I317333     C111565    Male   21     Shoes         3   1800.51
2    I127801     C266599    Male   20  Clothing         1    300.08

    payment_method invoice_date   shopping_mall
0     Credit Card     5/8/2022          Kanyon
1      Debit Card   12/12/2021  Forum Istanbul
2            Cash    9/11/2021       Metrocity
```

```
         payment_method invoice_date   shopping_mall
      0     Credit Card     5/8/2022          Kanyon
      1      Debit Card   12/12/2021  Forum Istanbul
      2            Cash    9/11/2021        Metrocity
```

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('/content/Iris.csv')
plt.bar(df['PetalLengthCm'],df['PetalWidthCm'])
plt.xlabel('Length')
plt.ylabel('Width')
plt.title('Bar Chart')
plt.show()
```
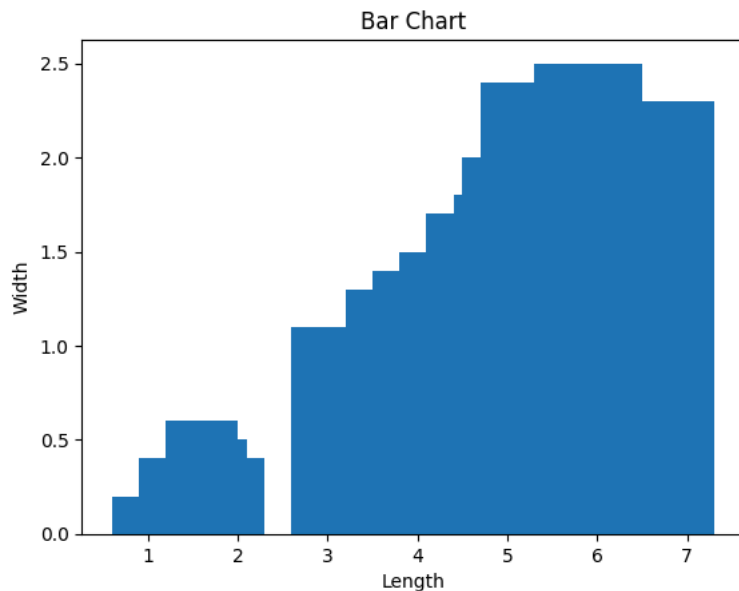


```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('/content/Iris.csv')
plt.plot(df["SepalWidthCm"], df["SepalLengthCm"])
plt.title("Line Chart")
plt.xlabel("Length")
plt.ylabel("Width")
plt.grid(True)
plt.show()
```
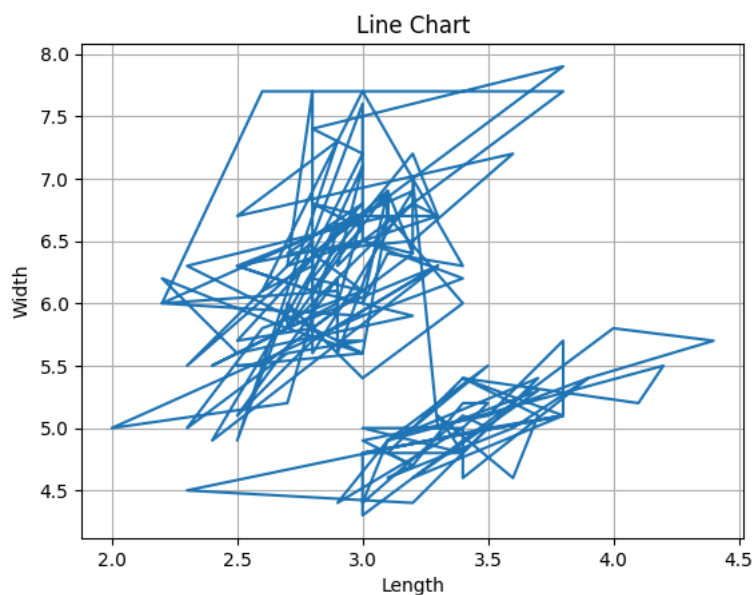
```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('/content/Iris.csv')
print("Displot")
sns.set_style('whitegrid')
sns.distplot(df['SepalWidthCm'], color ='red').set(title='Displot')
```

⤓ Displot
      <ipython-input-38-303bdbfd5689>:9: UserWarning:

      `distplot` is a deprecated function and will be removed in seaborn v0.14.0.

      Please adapt your code to use either `displot` (a figure-level function with
      similar flexibility) or `histplot` (an axes-level function for histograms).

   For a guide to updating your code to use the new functions, please see
   https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

        sns.distplot(df['SepalWidthCm'], color ='red').set(title='Displot')
      [Text(0.5, 1.0, 'Displot')]
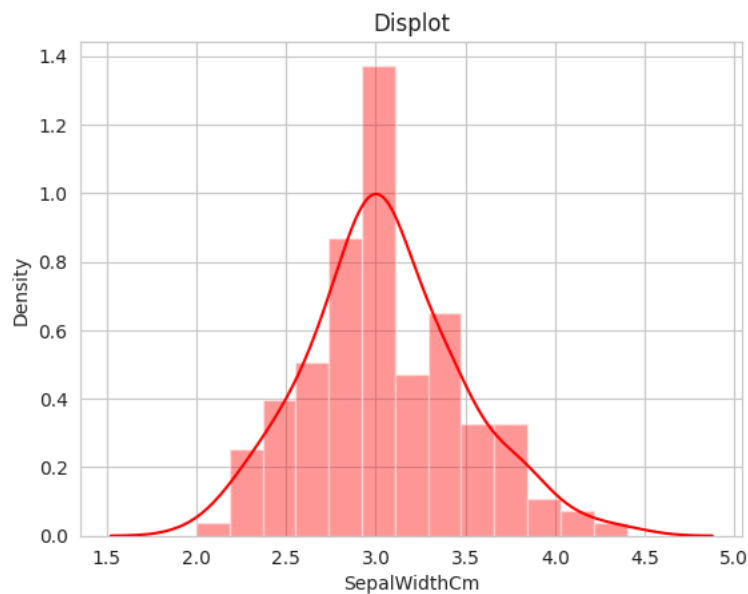

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv('/content/Iris.csv')
x=sns.pairplot(df)
x.fig.suptitle("My Pairplot")
plt.show()
```
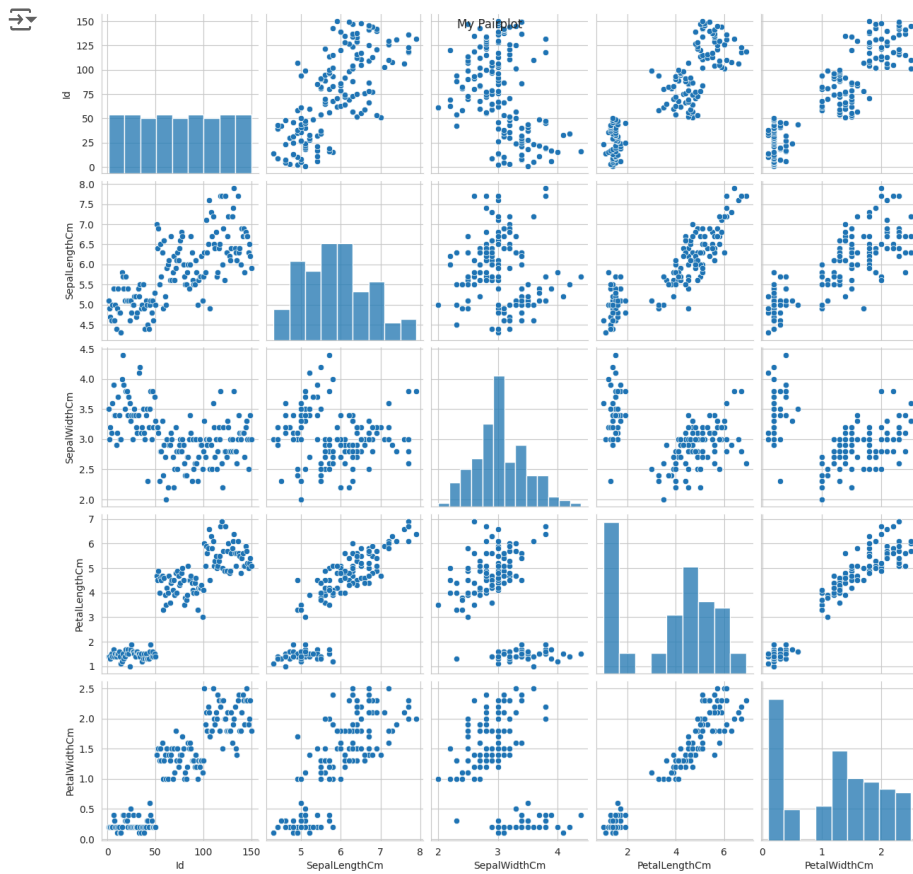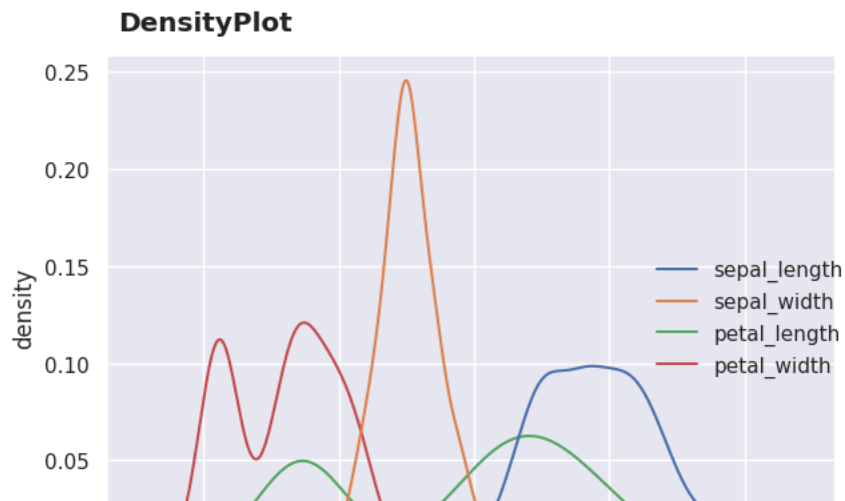
My Pairplot

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv('/content/Iris.csv')
plt.figure(figsize=(10, 6))
plt.hist(df['PetalLengthCm'], bins=20, color='blue')
plt.title('Histogram')
plt.xlabel('x-axis')
plt.ylabel('y-axis')
plt.show()
sns.histplot(df['SepalLengthCm'], label='Sepal Length')
plt.show()
sns.histplot(df['SepalWidthCm'], label='Sepal Width')
plt.show()
```
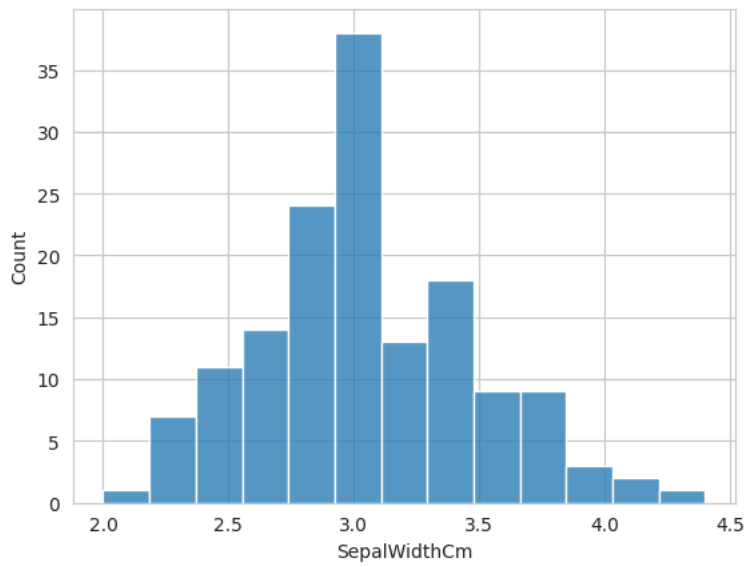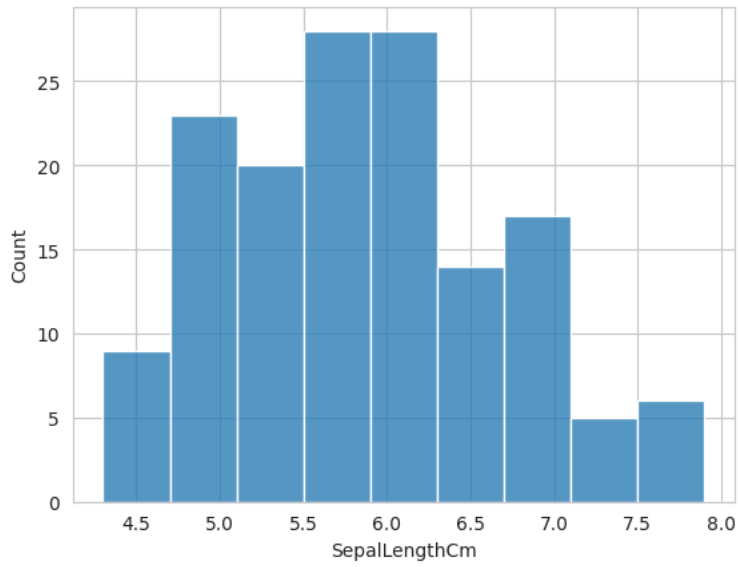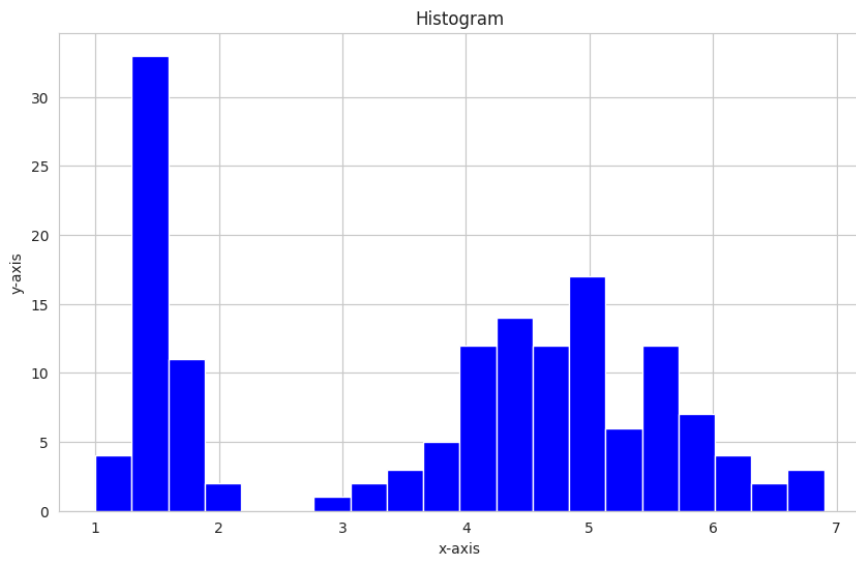
Histogram

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df = sns.load_dataset('iris')

sns.displot(df, kind="kde", color = 'black')
plt.suptitle("DensityPlot", x=0.149, y=0.96, ha='left', fontweight = 'bold')
plt.xlabel("x-axis")
plt.ylabel("density")
plt.tight_layout()
plt.show()
```



```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df = sns.load_dataset('iris')
df_numeric = df.select_dtypes(include=[int, float])
corr = df_numeric.corr()
plt.figure(figsize=(10, 8))
sns.heatmap(corr, annot=True, cmap='coolwarm')
plt.title('Heatmap of Iris Dataset')
plt.show()
```